FIMAD

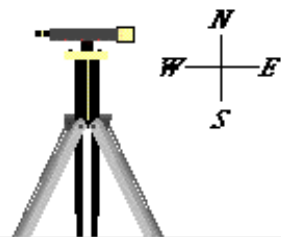## Facility for Information Management, Analysis & Display

Update on
Analytic Chemistry Data

Thom McVittie

Agenda

- ❖ **Overview of data feeds**
- ❖ **Non Standard Requests**
- ❖ **Standard Requests via EDD**
- ❖ **Data Review/Approval**
- ❖ **Revised Analytic Tables**
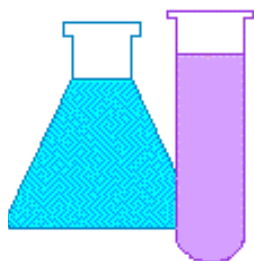- ❖ **Current Status & Time Line**
- ❖ **What's next?**

Major Types of Tabular Data
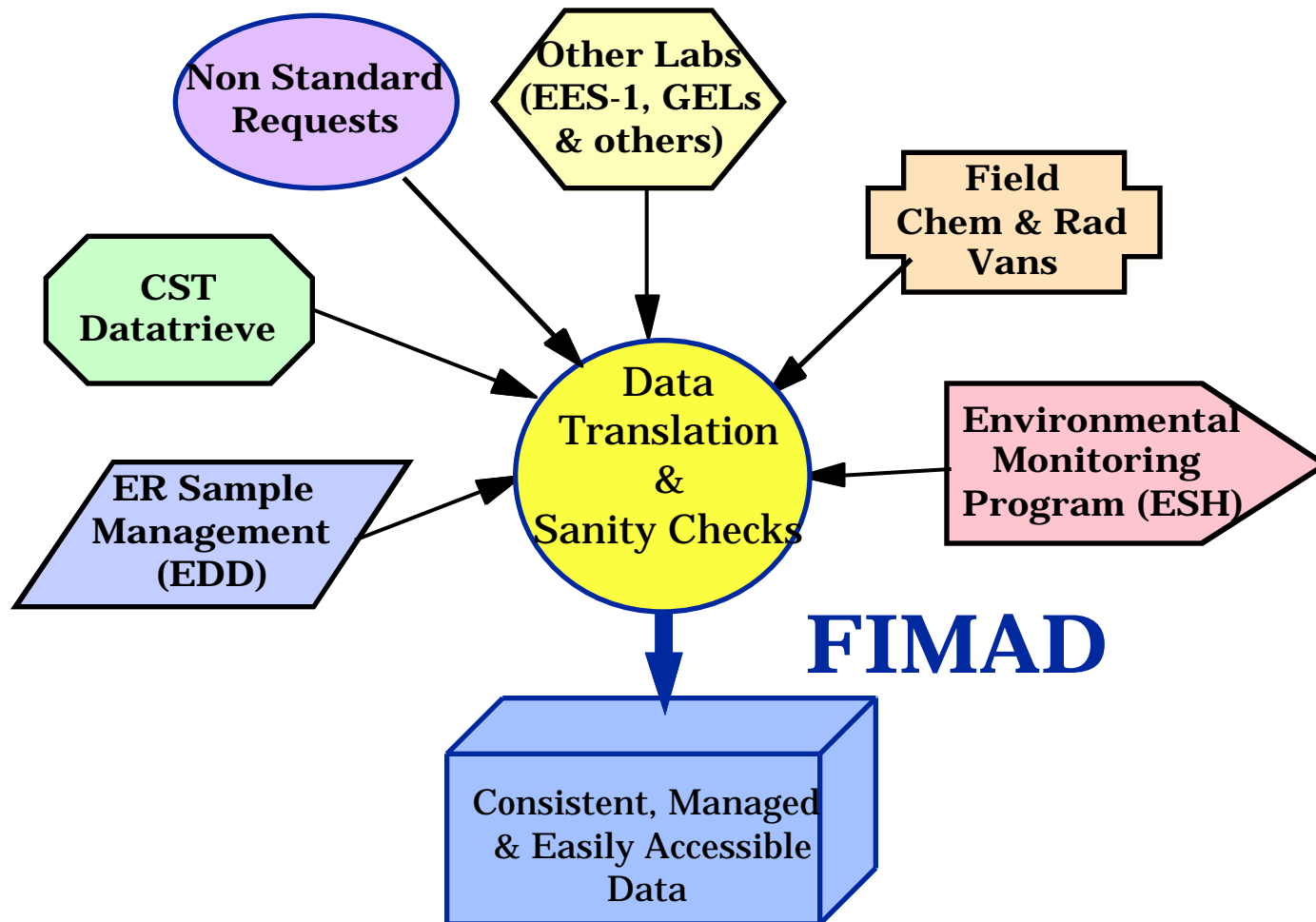
Location Data

Field Sample Info

Sample Results

SAL

Limits

Analytic Data Paths

Non Standard Requests

Other Labs (EES-1, GELs & others)

Field Chem & Rad Vans

CST Datatrieve

Data Translation & Sanity Checks

Environmental Monitoring Program (ESH)

ER Sample Management (EDD)

FIMAD

Consistent, Managed & Easily Accessible Data

Status of Data Feeds

❖ **CST**

❖ **GEL (Pilot Study)**

- data available in FIMAD in independent tables
- will integrate but still some issues to resolve

❖ **Chem Van/ Rad Van**

- Rad Van ready to transmit data in next three weeks
- Chem Van ??

❖ **Non Standard Requests**

- forms completed yesterday

❖ **SMO (Standard Requests)**

- almost done
- bulk of this presentation

Non Standard Analyses

Non Standard Requests → Manual Entry using Forms → FIMAD Stage Tables

❖ **Manual Entry using Forms**

  – on the spot consistency checks
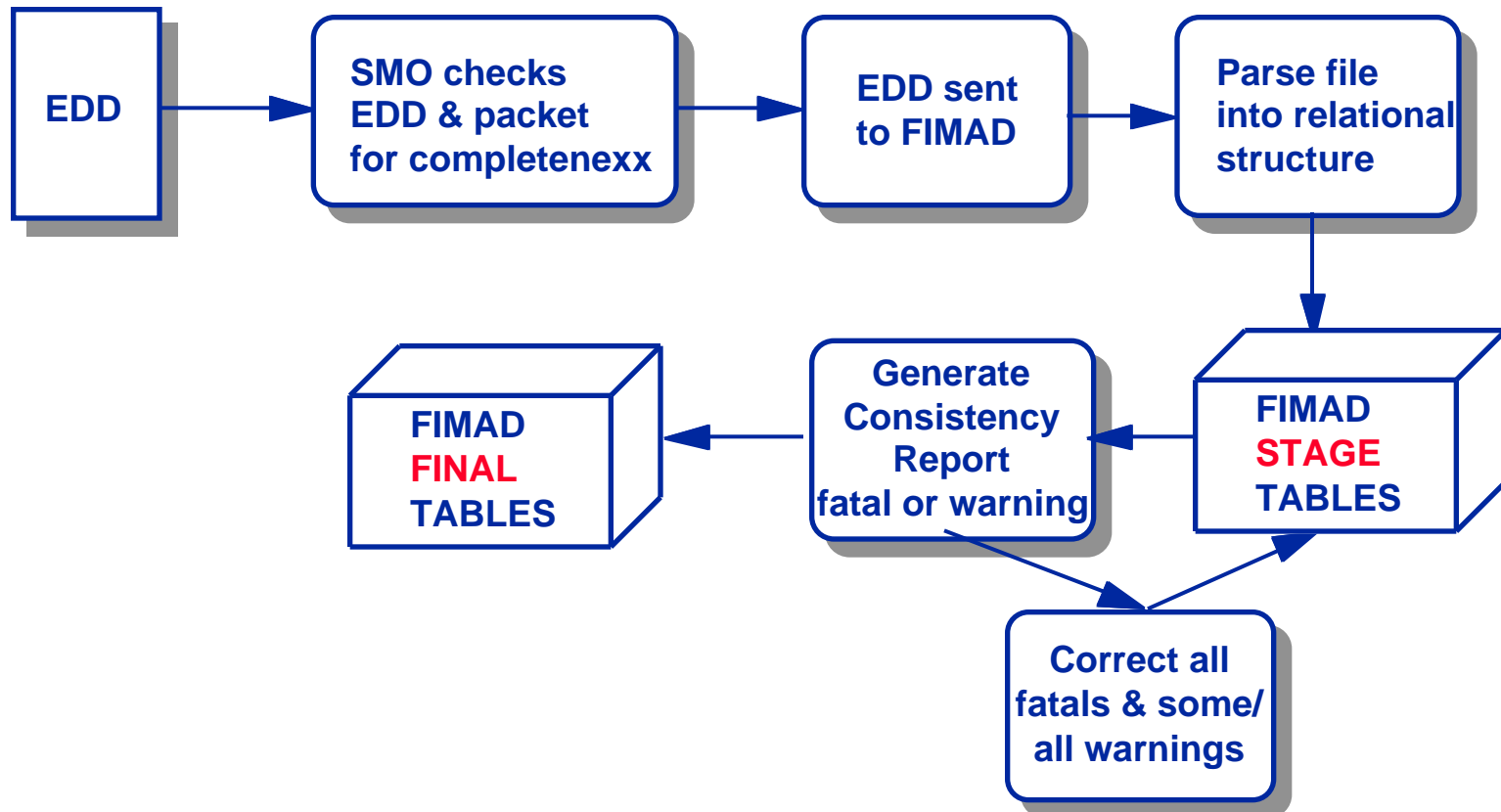
  – tracking of who did what

❖ **Forms are now complete**

  – Clovis Martinez

Overview of EDD

❖ **A different EDD for each type of standard analysis.**

❖ **Spreadsheet-like format**

- – each row contains a given type of data
- – each column represents the data for one cut/analysis.
- – an SDG may be in a single file or span multiple files.

❖ **Nine different Sections**

- –

EDD Process

EDD → SMO checks EDD & packet for completenexx → EDD sent to FIMAD → Parse file into relational structure → FIMAD STAGE TABLES → Generate Consistency Report fatal or warning → FIMAD FINAL TABLES

Correct all fatals & some/all warnings

Problems with the EDDs

❖ **Labs weren't given specific instructions on how to fill out.**

- – Analytical Method Code? EPA method or SMO order code or ????
- – Recovery Data -- where and in what format (85.3 vs .853)

❖ **They contained Errors**

- – Concentration rather than percent recovery on many surrogates

❖ **Missing key data fields**

❖ **Every EDD submission was different!**

The Fix (patch)

- ✓ **Review all EDDs for correctness and classify fields as Optional, Required, or don't care.**
- ✓ **Provide a data dictionary and instructions for the Labs.**
- ✓ **Bring the labs to LANL to introduce corrected EDD and get their comments/suggestions.**

- ❖ **Begin receiving data in corrected format.**

The Fix II

❖ **The data received via the original EDD is flawed.**

❖ **Automated fixes are done where possible**

– e.g., repair matrix, sample type, extract qualifiers, etc.

❖ **Manual "informed" corrections may be needed until the new EDD formats are used.**

– e.g., missing lab batch, missing unit of measure, bad sample id, etc.

❖ **Corrections via the SAME form as non-standard requests.**

Four Main Tables

❖ **Sample Request Header**

 – data concerning the request and the sample, e.g., LANL sample number, request number, suite, etc.

❖ **Analysis Header**

 – data specific to an analysis - e.g., date of prep, prep method, dilution factor, etc.
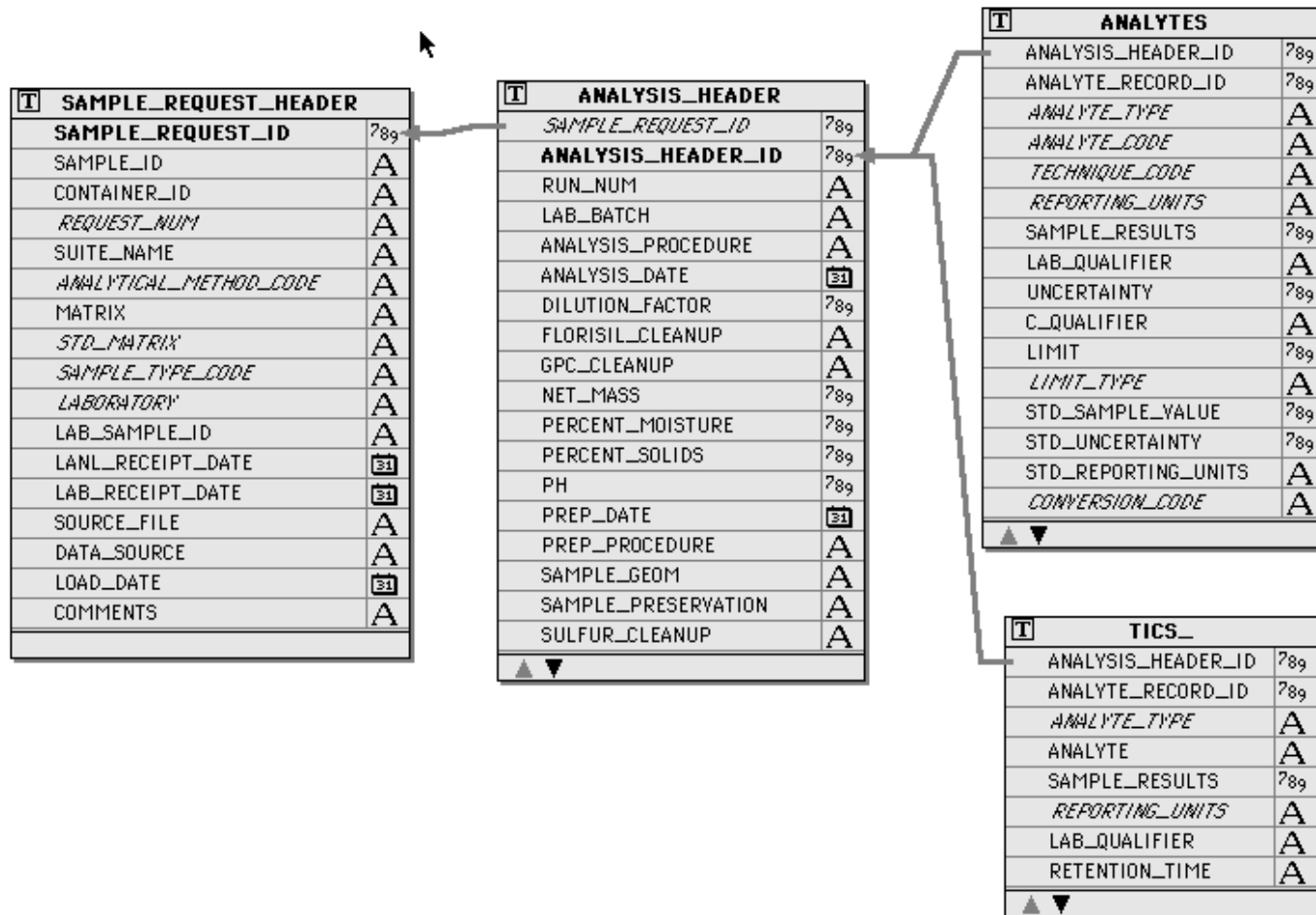
❖ **Analyte**

 – Analytic results

❖ **TICS**

 – Data concerning tentatively identified compounds.

Overview

Schema

❖ **Very Relational**

- data appears only once - e.g., sample number only in sample request header
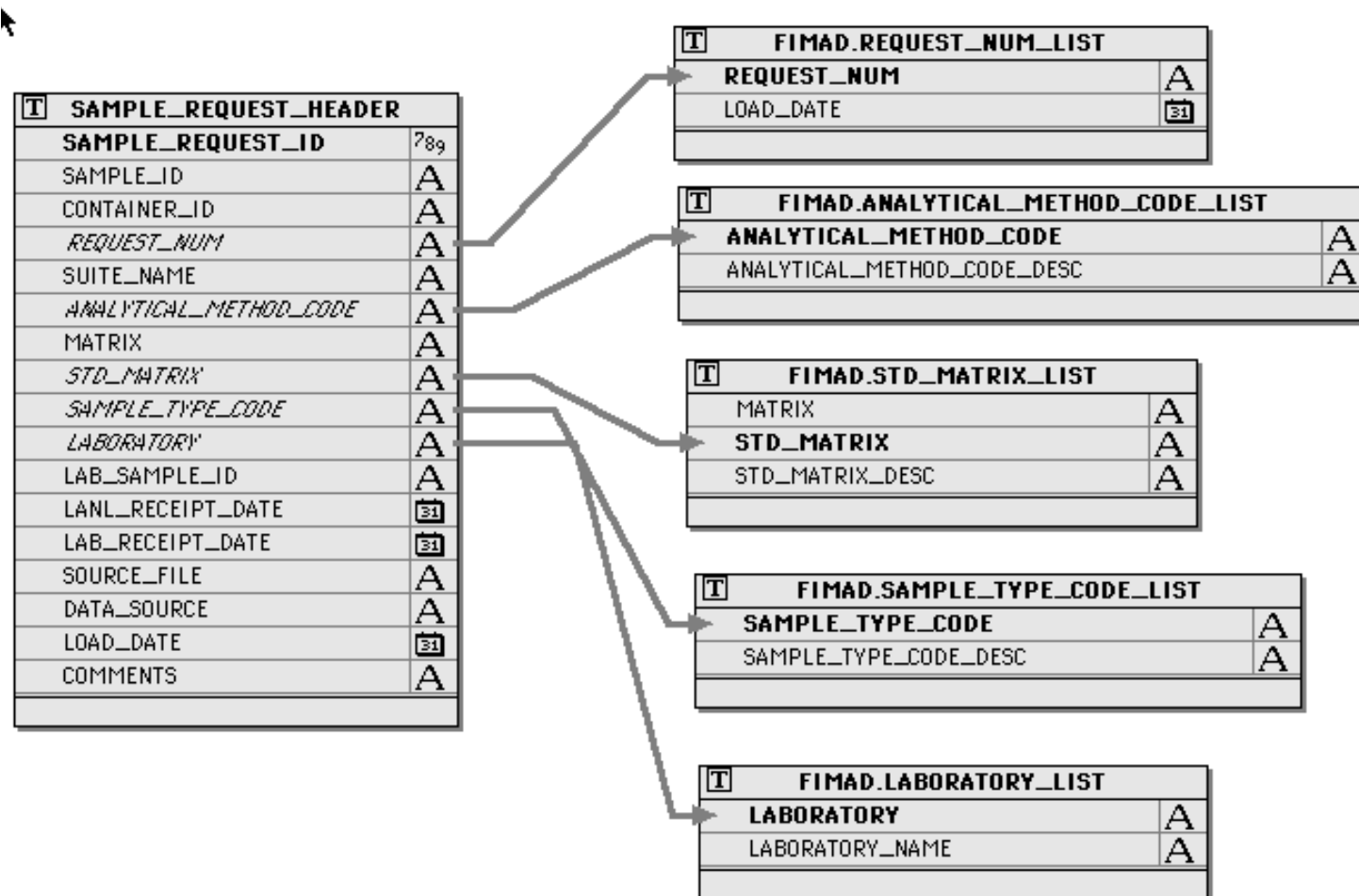- links provided between the various tables are all derived.

❖ **Why?**

- improved access speed (queries will complete faster)
- less storage space.
- easier to correct erroneous data

❖ **Too Complex ?**

- views will be provided which will make data look flat.
- tables "know" about the joins and in some applications will make them automatically.
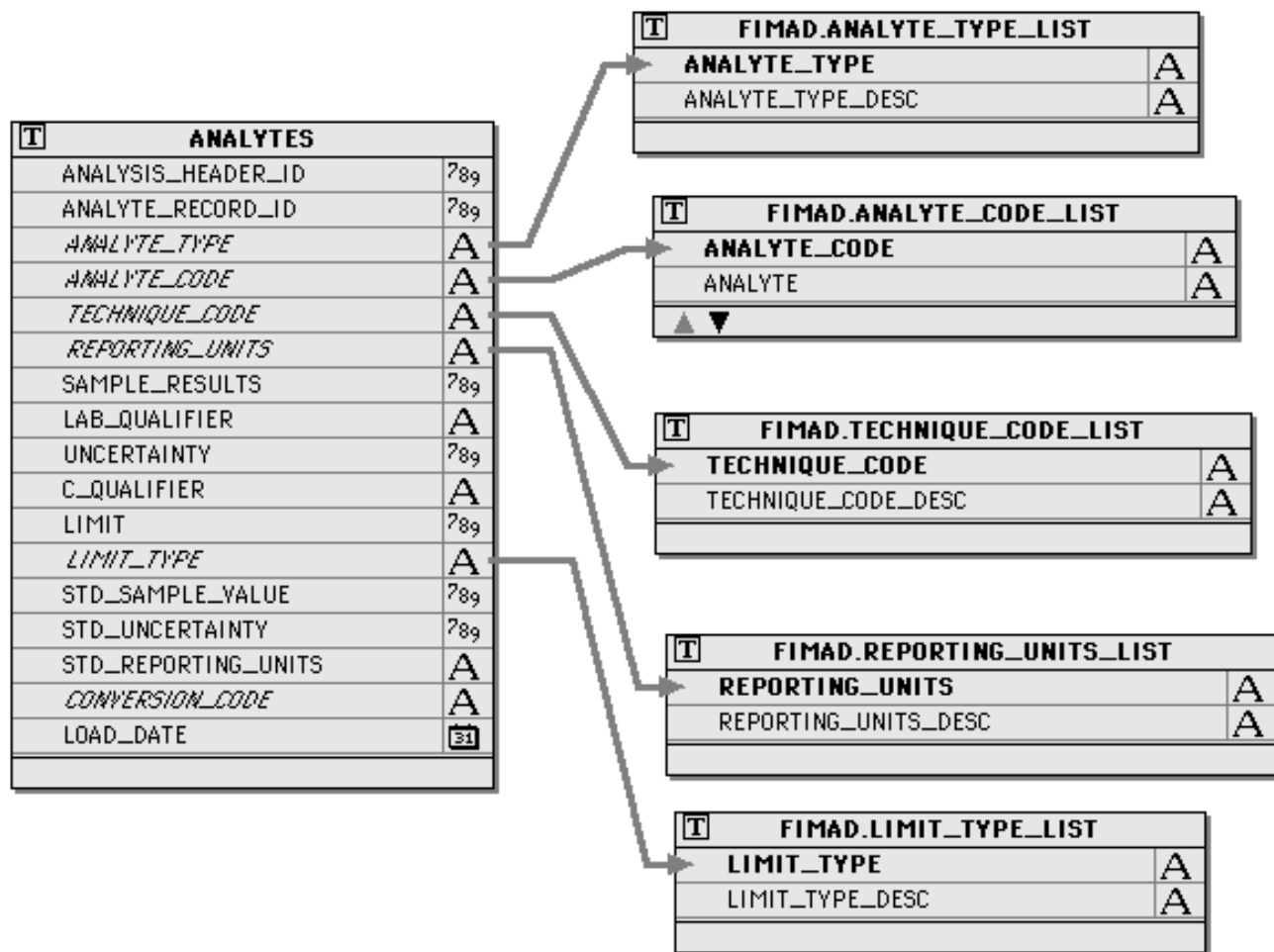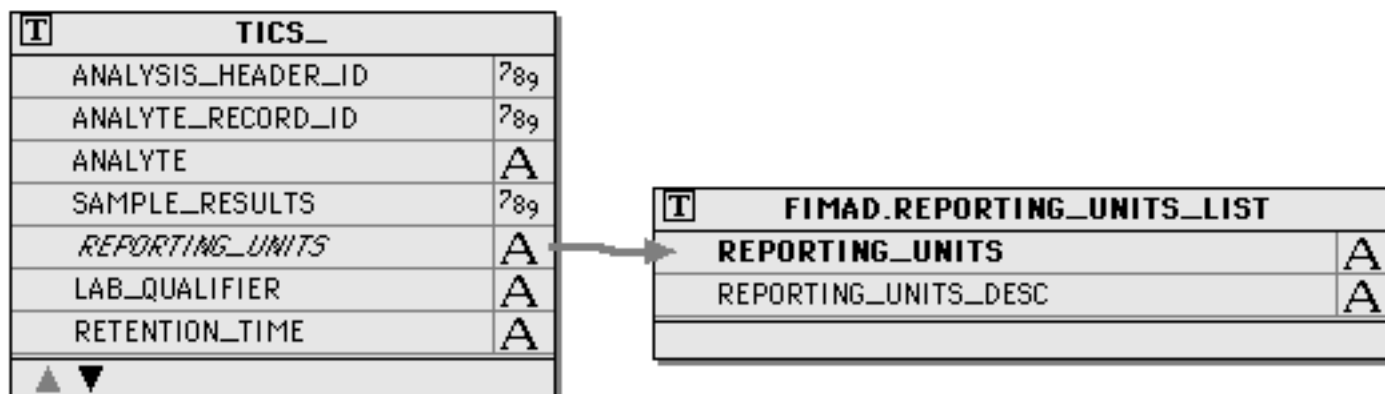
Sample Request Header

Analysis Header

ANALYSIS_HEADER

SAMPLE_REQUEST_ID	789
ANALYSIS_HEADER_ID	789
RUN_NUM	A
LAB_BATCH	A
ANALYSIS_PROCEDURE	A
ANALYSIS_DATE	31
DILUTION_FACTOR	789
FLORISIL_CLEANUP	A
GPC_CLEANUP	A
NET_MASS	789
PERCENT_MOISTURE	789
PERCENT_SOLIDS	789
PH	789
PREP_DATE	31
PREP_PROCEDURE	A
SAMPLE_GEOM	A
SAMPLE_PRESERVATION	A
SULFUR_CLEANUP	A
LOAD_DATE	31

Analyte

TICS

Example Query

Retrieve analytic data for sample 0121-95-0111

Select

        sample_id, sample_type_code, request_num,
        dilution_factor, analysis_procedure,
        analyte_code, sample_result, reporting_units,
        lab_qualifier

From

        sample_request_header a,
        analysis_header b,
        analytes c

Where

        a.analysis_header_id = b.analysis_header_id and
        b.analyte_record_id = c.analyte_record_id and
        sample_id = '0121-95-0111' ;

Example Query with View

Retrieve analytic data for sample 0121-95-0111

Select

> **sample_id, sample_type_code, request_num,
> dilution_factor, analysis_procedure,
> analyte_code, sample_result, reporting_units,
> lab_qualifier**

From

> **sample_view_fu01**

Where

> **sample_id = '0121-95-0111' ;**

2.5 Million Dollar Questions

Where the heck is my data?

Why is this taking so long?

Current EDD Status

Files received: 2,500

Duplicates   <300>   2,200

In Stage

Need Resolution

Dead

Ready to Process

data tracked in table edd_file_status

A few (more) Quirks

❖ **Dual Column data (sometimes)**

❖ **Reporting values detected at less than the reporting limit.**
- Rads: MDAs and MDCs
- Nothing similar for Organics/Inorganics

❖ **Surrogate recoveries**
- Watch out!
- Percent recovered vs amount spiked

❖ **Inconsistent reporting**
- LCS/MS recoveries

The Schedule

❖ **Non-standard analysis entry forms available**

 – end of this week

❖ **Data correction forms available**

 – end of next week

❖ **Finish EDD backlog**

 – by end of February

❖ **Convert CST and GEL data to same format**

 – need significant input from chem team and others

 – tentative end of March

Where to Next?

❖ **LANL EDD not sufficient**

  – inconsistent and cumbersome for labs

  – doesn't supply all fields needed for basic electronic V/V
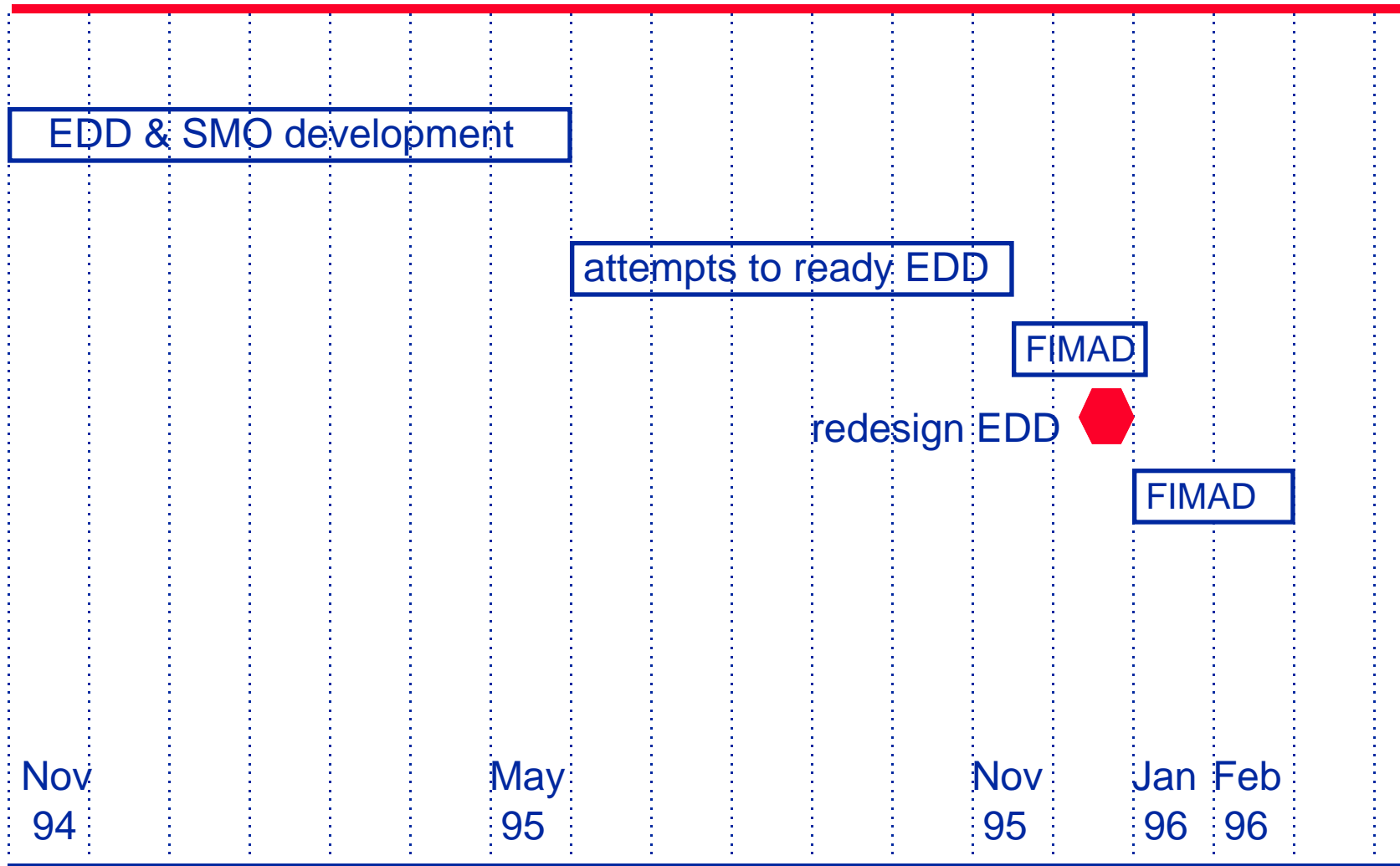
❖ **DEEMS best approach**

  – well defined

  – relatively easy to write/read

  – wider support for tool sets

  – labs interested in having a smaller set of EDDs.

❖ **Potential Pitfalls with DEEMS**

  – DOE complex needs to standardized on a set of EDDs based on DEEMS.

  – significant effort to retool systems.

EDD Timeline

EDD & SMO development

attempts to ready EDD

FIMAD

redesign EDD

FIMAD

Nov	May	Nov	Jan	Feb
94	95	95	96	96

Accessing Data

❖ **Detailed data definitions**

  – http://ees.lanl.gov/fimad/database

❖ **Various methods to access all tables:**

  – FIMAD supplied software

    » Browser

    » SAS & S-Plus

    » Arcview

    » SQLPlus

  – SQL connections

    » connections to MS Access & Excel

    » connections to Fox Pro, Dbase

    » lots of others